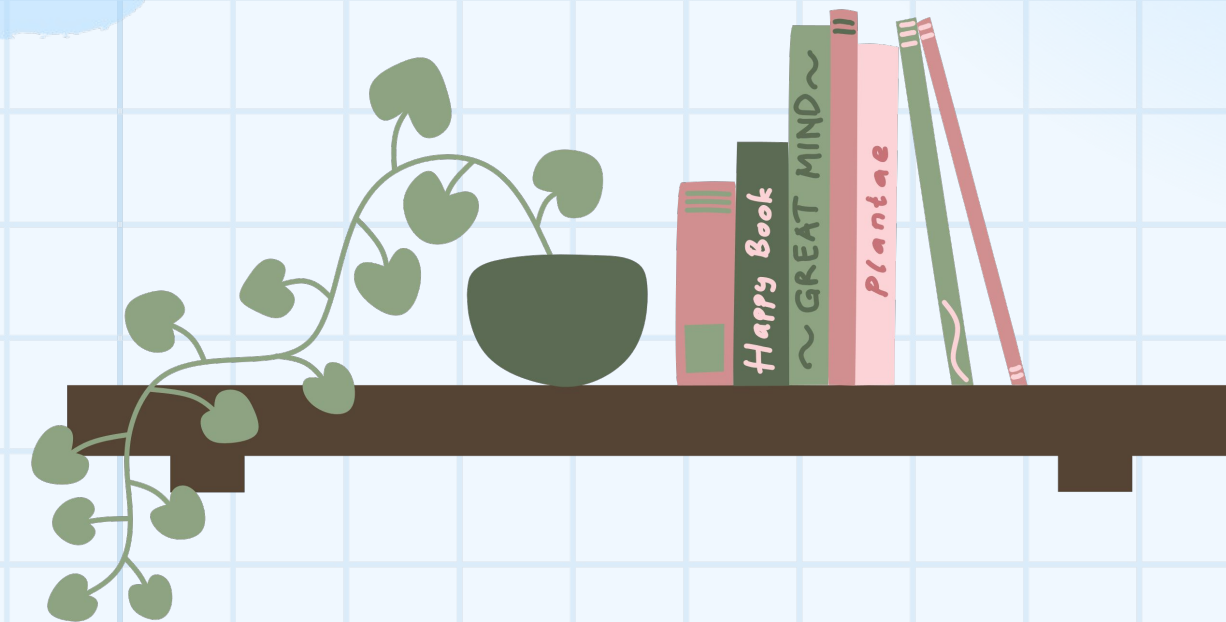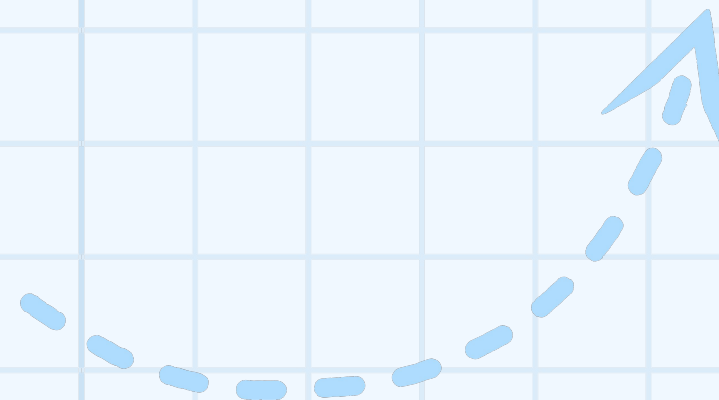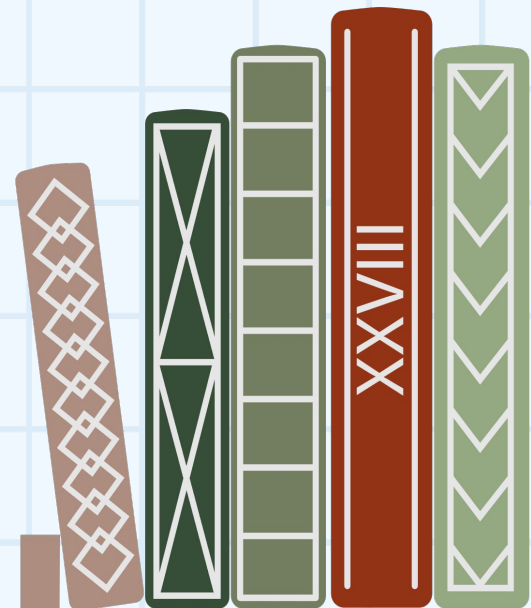# BS./BSC.IN
Applied AI and Data Science

Basics of Data Analytics

Let's dive into and learn:

1   NoSQL Databases

Non-Relational Databases

- Non-relational databases emerged in response to the volume, variety, and speed at which data is being generated today

- Built for speed, flexibility, and scale, non-relational databases made it possible to store data in a schema-less or free-form fashion.

- NoSQL is widely used for processing such data

No SQL

- NOSQL (not only SQL) is a non-relational database design that provides flexible schemas for storage and retrieval of data

- Recently become popular in the era of dynamic, diverse, high-volume data

- Popular due to their attributes around scale, performance

- Easy to use

No SQL

- Usually built for data-specific models

- Has flexible schemas to support that allow the creation and support many applications

- Don't use the neat row and column structure

- Typically, don't use SQL to query data

- Allows data to be stored in a schema less format

No SQL

- Can be of the following types

  - Key –Value store

  - Document based store

  - Column based store

  - Graph based store

Key Value Based Store

- Data stored as a collection of key-value pairs

- Key is a characteristic or attribute of the data

- Key is a unique identifier

- Values are all the information related to each key

- Supports a wide variety of formats of data for both keys and values

- Useful in storing user session data, user preference, real-time recommendations, targeted advertising

Disadvantages of Key Value

- Not appropriate for querying data on a specific value

- Not appropriate if you need to get the relationship between data values

- Examples are Redis , Memcached

Document Based Store

- Document databases store each entry of data within a single document

- Can do ad-hoc queries and analytics over a set of documents

- Usually used in e-commerce platforms, invoices, medical records

Disadvantages of Document Based

- Hard to run search queries

- May be difficult to retrieve data based on specific queries

- E.g. MongoDB

Column Based Store

- Store data in cells grouped as columns of data

- Columns are grouped together into column family based on some

  logical condition

- E.g. purchases and payment modes can be a column family

- Cells corresponding to a column are stored as a continuous entry

- Makes access and search faster

- Great for storing time-series data, weather data, IoT data

Disadvantages of Column Based

- Hard to run complex queries

- Difficult to change the querying pattern

- Not flexible querying

- E.g. Cassandra

Graph Based Store

- Graph based data storage uses a graphical model to represent and

  store data

- Great for data that are connected to each other in some way.

- Can be used for visualizing and analyzing as well

Graph Based Store

- Each circle is a node and it contains the data

- The lines are the connections and represent the relationships

- Graph based store is great for social networks, real-time product recommendations, network diagrams
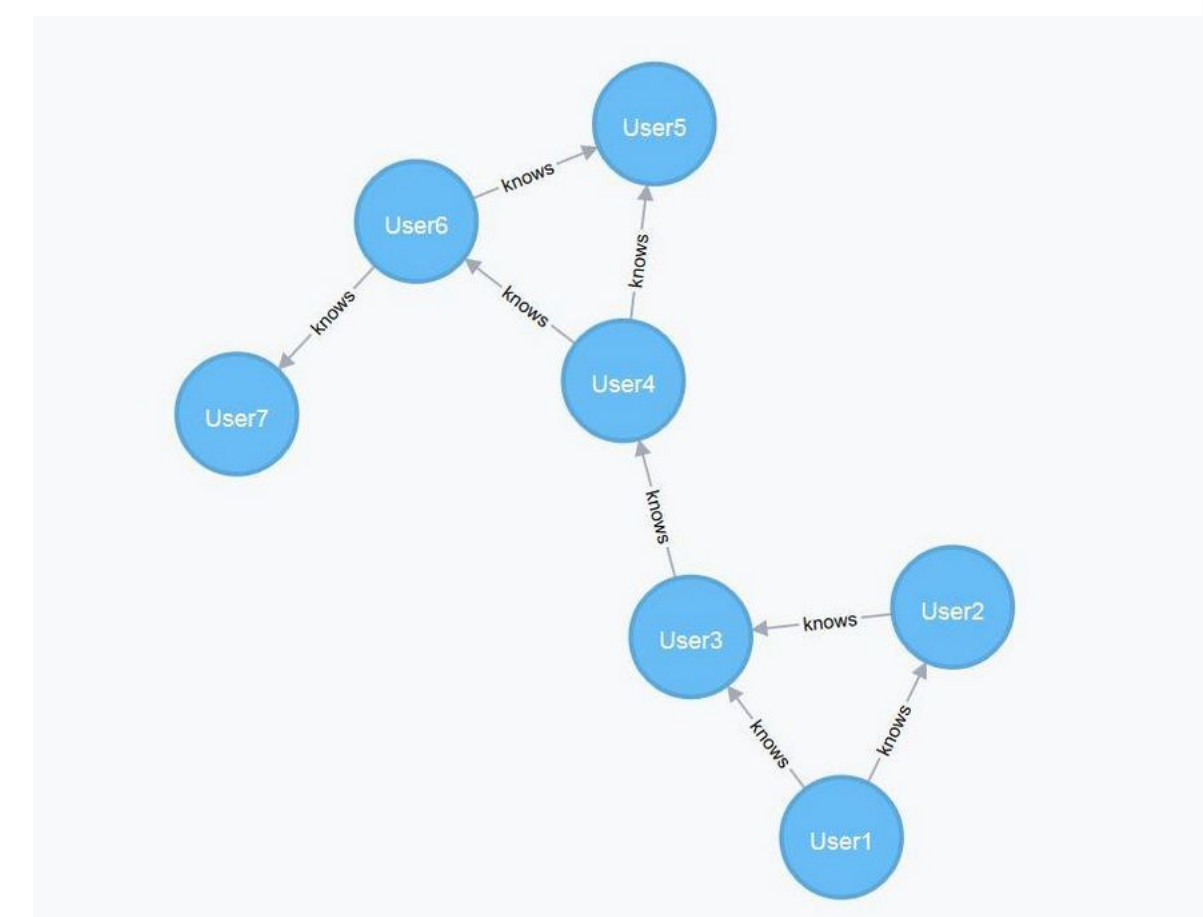
Graph Based Store

- Can be quite versatile

- Difficult to run large volumes of data

- Not optimized for large volumes analytics

  queries

- E.g. Neo4J, CosmosDB

Advantages of NOSQL

- Ability to store and handle large volumes of structured,

  semi-structured and unstructured data

- Schema agnostic

Recap

Different types of NOSQL

databases